

<p>(51) International Patent Classification ⁶ : C12N 15/29, C07K 14/415, C12N 5/10, A01H 5/00, C07K 16/16</p>	<p>A2</p>	<p>(11) International Publication Number: WO 99/58681</p>
<p>(21) International Application Number: PCT/EP99/03158</p>		<p>(43) International Publication Date: 18 November 1999 (18.11.99)</p>
<p>(22) International Filing Date: 7 May 1999 (07.05.99)</p>		
<p>(30) Priority Data: P 9800975 8 May 1998 (08.05.98) ES P 9800981 11 May 1998 (11.05.98) ES</p>		
<p>(71) Applicant (for all designated States except US): CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES/ES); Serrano, 113, E-28006 Madrid (ES).</p>		<p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p>
<p>(72) Inventors; and</p>		<p>Published</p>
<p>(75) Inventors/Applicants (for US only): GUTIERREZ-ARMENTA, Crisanto [ES/ES]; Centro de Biología Molecular, CSIC-UAM, E-28049 Madrid (ES). XIE, Qi [ES/ES]; Centro de Biología Molecular, CSIC-UAM, E-28049 Madrid (ES). RAMIREZ PARRA, Elena [ES/ES]; Centro de Biología Molecular, CSIC-UAM, E-28049 Madrid (ES).</p>		<p>Without international search report and to be republished upon receipt of that report.</p>
<p>(74) Agent: UNGRIA, Javier, Ungria Patentes y Marcas, S.A., Avenida Ramón y Cajal, 78, E-28043 Madrid (ES).</p>		

(54) Title: TRANSGENIC PLANT CELLS EXPRESSING A RECOMBINANT PLANT E2F PEPTIDE

[illegible]

(57) Abstract

A method of controlling plant growth and/or cellular DNA replication and/or cell cycle progression, differentiation and development comprising increasing or decreasing E2F activity in a plant cell.

TRANSGENIC PLANT CELLS EXPRESSING A RECOMBINANT PLANT E2F PEPTIDE

The present invention relates to nucleic acid sequences encoding plant E2F proteins or functional variants thereof, including peptides, and the use of said sequences for controlling the plant cell cycle stage and or its body architecture. The invention also provides plant E2F proteins and peptides useful in producing antibodies, and provides nucleic acids suitable for use in detection and amplification of plant E2F peptides and proteins. Further provided are transgenic plants, plant parts and plant cells overproducing or underproducing E2F protein and parts thereof involved in the mechanism of transition of plant cells from G1 to S phase in the cell cycle. Such plants, parts and cells may over or underproduce other proteins by virtue of being caused to increase or decrease the amount of time they spend in G1 or S phase.

Cell cycle progression is the result of a complex and highly regulated network. Crucial for the correct passage of the cell through the different cell cycle stages is the strict regulation of the transcriptional activity of certain genes, e. g., S-phase specific genes (reviewed in Nevins, 1992; Helin, 1998).

In mammalian cells, the E2F family of transcription factors play this pivotal role in transcriptional regulation at the G1/S transition. Their concerted action is thought to modulate the expression of cell cycle regulatory genes such as cdc2, cyclins A and E, Rb, p107 and E2F-1, and genes involved in DNA metabolism, such as the dihydrofolate reductase, thymidine kinase, thymidylate synthase, DNA polymerase α , ORC1 and CDC6 (reviewed in Nevins, 1992; Helin, 1998). E2F activity on gene expression is mediated by the retinoblastoma (Rb) tumor suppressor protein as well as by its related p107 and p130 proteins through the formation of complexes between the different E2F members and pocket proteins (reviewed in Weinberg, 1995). In this way, for example, Rb is targeted to E2F-responsive gene promoters and inhibits transcription through interaction with adjacent factors, as recently shown for histone deacetylase (Brehm et al., 1998; Magnaghi-Jaulin et al., 1998).

In other systems, such as plants, which have unique properties in terms of cell growth and plasticity, body organization and development, the factors involved in cell

cycle regulation, in particular at the G1/S transition, and their mechanism of action are significantly less understood. However, the available data indicates that a strict control of gene expression, linked to and responsible for cell cycle progression, also exists in plant cells whereby some genes are known to be expressed at specific stages throughout the cell cycle (reviewed in Staiger and Doonan, 1993; Doonan and Fobert, 1997). For example, B-type cyclins accumulate in G2 and M phases (Ferreira et al., 1994; Fobert et al., 1994; Kouchi et al., 1995; Ito et al., 1997; Ito et al., 1998) while the ribonucleotide reductase and the histone genes mRNAs appear to be S-phase specific (Philipps et al., 1995; Shen and Gigot, 1997). Thus, the existence of S-phase specific transcription factors is possible in plant cells, but their molecular nature is not known yet. In particular, whether they have any structural and/or functional similarity to the animal E2F family of transcription factors is one of the important questions that still needs to be answered. In addition, it is known that the activity of S-phase specific protein kinases increases during early stages of endosperm development (Grafi and Larkins, 1995).

The first indications that a Rb-like pathway could regulate the G1/S transition in plant cells came after the isolation of three different D-type cyclins in plants (Soni et al., 1995; Dahl et al., 1995) and the observation that a protein from a plant DNA virus, whose replication depends on host functions, can associate with human Rb-related proteins (Xie et al., 1995). Later, plant cDNAs encoding proteins with a conserved A/B pocket domain were isolated (Xie et al., 1996; Grafi et al., 1996; Ach et al., 1997a).

Plant Rb-like protein has some features in common with its human counterpart, including the presence of a residue homologous to C607 of human Rb required for its activity and its ability to interact with the three plant D-type cyclins in a LXCXE-dependent manner (Huntley et al., 1998). Furthermore, quite interestingly, when plant Rb is expressed in human cells, it is able to repress an E2F-responsive promoter (Huntley et al., 1998). Altogether, these studies predict the existence of S-phase specific transcription factors (STF) in plant cells (Xie et al., 1995), perhaps related to the E2F family of transcription factors found in animal cells. However, the identification of E2F-like transcription factors in plants has been elusive since studies

using heterologous probes derived from human E2F cDNA clones have been unsuccessful.

The present inventors have now isolated, cloned and characterized cDNA encoding a plant protein which interacts with plant Rb in the yeast two-hybrid system. They have established that this cDNA clone encodes a plant E2F family member (TmE2F) with amino acid homology to animal E2F proteins. The inventors have further determined that, surprisingly, plants appear to contain a single E2F member with a domain organisation similar to that of human E2F, including a highly conserved DNA binding domain, a less conserved dimerization domain and relatively unrelated transactivation and Rb-binding domains. Interestingly, its Rb-binding domain contains amino acid residues different from those found in animal E2F but showing conservation of their hydrophobic or charged properties.

With respect to the present specification and claims, the following technical terms are used in accordance with the definitions below.

A "functional variant" of a peptide or protein is a polypeptide the amino acid sequence of which can be derived from the amino acid sequence of the original peptide or protein by the substitution, deletion and/or addition of one or more amino acid residue in a way that, in spite of the change in the amino acid sequence, the functional variant retains at least a part of at least one of the biological activities of the original protein that is detectable for a person skilled in the art. A functional variant is generally at least 50% homologous, advantageously at least 70% homologous and even more advantageously at least 90% homologous to the protein from which it can be derived. Preferably the amino acid sequence of the functional variant is 50% identical, more preferably 70% identical and most preferably 90% identical to the peptide or protein. Any functional part of a protein or a variant thereof is also termed functional variant.

Algorithms and software suitable for use in aligning amino acid or nucleotide sequences for comparison and calculation of sequence homology or identity will be known to those skilled in the art. Significant examples of such tools are the Pearson and Lipman search based FAST and BLAST programs. Details of these may be found in Altschul et al (1990), J. Mol. Biol. 215: 403-10; Lipman D J and Pearson W R

(1985) Science 227, p1435-41. Publically available details of BLAST may be found on the internet at 'http://www.ncbi.nlm.nih.gov/BLAST/blast-help.html'. Thus such homology and identity percentages can be ascertained using commercially or publically available software packages incorporating, for example, FASTA and BLASTn software or by computer servers on the internet. Examples of the former are the GCG program package (Devereux et al Nucleic Acids Research (1984) 12 (1): 387) and the Bestfit program (Wisconsin Sequence Analysis Package, eg. Version 8 for Unix or IBM equivalent, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711) which uses the local homology algorithm of Smith and Waterman, Advances in Mathematics 2:482-489 (1981). Many international units, eg. Genbank (see http://www.ncbi.nlm.nih.gov/BLAST) and EMBL: (see http://www.embl-heidelberg.de/Blast2), offer internet services.

By the term identity is meant that the stated percentage of the claimed amino acid sequence or base sequence is to be found in the reference sequence in the same relative positions when the sequences are optimally aligned, notwithstanding the fact that the sequences may have deletions or additions in certain positions requiring introduction of gaps to allow alignment of the highest percentage of amino acids or bases. Preferably the sequence are aligned by using 20 or less gaps, ie. the total number of gaps introduced into the two sequences when added together is 20 or less, more preferably 10 or less. The length of such gaps is not of particular importance as long as one or other of the two defined E2F activities is retained but generally will be no more than 50, and preferably no more than 10 amino acids, or 150 and preferably no more than 30 bases.

Parameters used in with software packages and internet servers should be applied with the appropriate sequence lengths and aforesaid gap characteristics in mind. Alignment strategies are discussed further in WO 98/40483 on pages 39 to 41, which document is incorporated herein by reference

Convenient parameters for BLAST searches are the default values, ie. for EMBL Advanced Blast2: Blastp Matrix BLOSUMS 62, Filter default, Echofilter X, Expect 10, Cutoff default, Strand both, Descriptions 50, Alignments 50. For BLASTn

defaults are again preferably used. GCG Wisconsin Package defaults are Gap Weight 12, Length weight 4. FASTDB parameters used for a further preferred method of homology calculation are mismatch penalty = 1.00, gap penalty = 1.00, gap size penalty = 0.33 and joining penalty = 30.0.

5 The term "overproducing" is used herein in the most general sense possible. A special type of molecule, usually a polypeptide or an RNA, is said to be "overproduced" in a cell if it is produced at a level significantly and detectably higher (e.g. 20% higher) than natural level. Overproduction of a molecule in a cell can be achieved via both traditional mutation and selection techniques and genetic
10 manipulation methods.

 The term "ectopic expression" is used herein to designate a special realisation of overproduction in the sense that, for example, an ectopically expressed peptide or protein is produced at a spatial point of a plant where it is naturally not at all (or not detectably) expressed, that is, said peptide or protein is overproduced at said point.
15 Particularly preferred ectopic expression is that which only reaches functional levels in a selected tissue and does not do so throughout the plant. This preferred ectopic expression is in contrast to constitutive expression.

 The term 'underproducing' is intended to cover production of polypeptide or mRNA at a level significantly lower than the natural level (eg. 20% or more lower),
20 particularly to undetectable levels.

 In a first aspect of the present invention there is provided a method of controlling plant growth and/or cellular DNA replication and/or cell cycle progression, differentiation and development comprising increasing or decreasing E2F activity in a plant cell through expression of a recombinant E2F peptide or
25 protein in that cell.

 Preferably the method is characterised in that the plant E2F activity comprises one or both of (i) the ability to bind plant Retinoblastoma protein and (ii) the ability to bind to E2F transcription factor binding sites in plant DNA. This may include steps of altering the plant E2F protein level, subcellular localisation, DNA-binding activity,
30 the protein-protein binding activity, transactivation properties, and/or the E2F-Rb-

binding activity. The plant E2F may be modified alone and/or in combination with a modification of the levels or activity of plant Rb.

The ability to bind to the E2F transcription factor binding sites in plant DNA need not necessarily lead to transcription activation. Inhibition of such activation can
5 also be provided using the present invention.

Particularly the method may be used to alter plant cell or organ shape, and it may alter cell proliferation characteristics such as to increase plant cell or plant organ size. The method may also increase or decrease expression of other proteins.

In a second aspect the present invention provides an isolated, enriched, cell
10 free and/or recombinantly produced protein or peptide, capable of altering E2F activity in a plant cell, characterised in that it has one or both E2F activities in plants selected from (i) the ability to bind plant Retinoblastoma protein and (ii) the ability to bind to E2F transcription factor binding sites in plant DNA

wherein the protein or peptide comprises one or both amino acid sequences
15 selected from the following domains of SEQ ID No 6:

(a) Tyr Xaa Xaa Xaa Xaa Xaa Xaa Asp Xaa Xaa Xaa Xaa Asp Met Trp Glu
and

(b) Gln Lys Arg Arg Ile Tyr Asp Ile Thr Asn Val Leu Glu Gly Ile Xaa Leu Ile
Glu Lys Xaa Xaa Lys Asn Xaa Ile Arg Trp

20 provided that where the peptide or protein comprises only domain (b) it comprises a sequence corresponding to at least 30% of the length of the contiguous sequence of amino acids 1-406 of SEQ ID No 6 or functional variants thereof.

More preferably the peptide or protein comprises at least 50% of the contiguous sequence and still more preferably at least 70% thereof. Most preferably
25 the peptide or protein comprises SEQ ID No 6 or a functional variant thereof. Preferred variants are those in which the domain (a) has been deleted or in which it is inactivated, eg. by Site directed mutagenesis.

Thus particularly preferred peptides or proteins of the invention are characterised in that they are of SEQ ID No 6 or variant but modified such that the
30 amino acid sequence SEQ ID No 2 is mutated such that its ability to bind Rb protein is reduced from that of the native sequence of SEQ ID No 2 or abolished completely

therefrom, whereby the peptide is capable of acting as an E2F protein without being restricted by Rb binding

It is particularly preferred that peptides or proteins of SEQ ID No 6 or functional variants thereof are provided that do not have the transinducing properties of the protein of SEQ ID No 6, these preferably having mutations or deletions or
5 insertions in the transinducing domain of SEQ ID No 6 in the C-terminal.

Preferred peptides or proteins of the invention are further characterised in that they comprises a sequence

(c) Leu Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Glu Xaa Xaa Xaa Leu Xaa
10 Xaa Xaa Glu Xaa Xaa Leu Asp

For some purposes it will be convenient to provide peptides or proteins of reduced length, for example 16 to 300, more preferably from 16 to 100 amino acids.

Further preferred peptides or proteins are characterised in that they comprise an amino acid sequence of SEQ ID No 2 or a functional variant thereof.

15 Still more preferred are peptides or proteins of the invention that are characterised in that they further comprises a sequence of SEQ ID No 7, that being of sequence

Arg Thr Gln Leu Lys Arg Lys Ala Thr Arg Glu Glu

or a functional variant thereof having functional activity as a nuclear
20 localisation signal (NLS).

Useful variants of such proteins however are those in which the NLS of SEQ ID No 7 is modified, eg. by site directed mutagenesis, eg using PCR, such that the peptide does not localise in the nucleus.

Further useful variants of the peptides or proteins of the invention are
25 characterised in that they comprise a plant E2F DNA binding domain being of sequence of amino acid residues 146-206 of the plant E2F of SEQ ID No 6 or a functional variant thereof.

Particularly preferred target E2F binding domains in plant DNA are of sequence TTT(C/G)(C/G)(C/G)(C/G)(C/G), particularly TTT(C/G)(C/G)CG(C/G).

30 In the case of an isolated, enriched, cell free and/or recombinantly produced peptide or protein comprising SEQ ID No 4

Tyr Xaa Xaa Xaa Xaa Xaa Xaa Asp Xaa Xaa Xaa Xaa Asp Met Trp Glu

or a functional variant thereof which lack other essential E2F peptide or protein regions, eg where it is a peptide of 16 to 100, more preferably 16 to 30 amino acids, it may be used to bind Rb and thus increase the effect of native E2F.

5 More preferably the peptide consists of SEQ ID No 4 or a functional variant thereof.

In some preferred forms the peptide is of SEQ ID No 2 but is modified such that the amino acid sequence SEQ ID No 4 is mutated such that its ability to bind Rb protein, eg. plant Rb protein, is increased or reduced from that of the native sequence
10 of SEQ ID No 4 or abolished completely therefrom. Particularly the peptide or protein is capable of acting as an E2F DNA binding, and optionally transcription activating, protein without being restricted by Rb binding. Such activity can then be more closely controlled using tissue specific or chemically inducible promoters

A third aspect of the present invention provides isolated, enriched, cell free
15 and/or recombinant nucleic acid comprising a sequence encoding for expression of a peptide as described in the first aspect of the invention. Preferred nucleic acids comprise DNA of less than 4,000 basepairs. Preferred nucleic acids comprise only one peptide or protein encoding DNA sequence, optionally together with a reporter gene.

20 Preferably the nucleic acid is that encoding for a plant E2F or a functional variant thereof including SEQ ID No 3, eg. being that of SEQ ID No. 1. Preferred nucleic acid comprises DNA or RNA of SEQ ID No 5 wherein when the nucleic acid is RNA the base T is substituted by U.

A nucleic acid of SEQ ID No 5 has been deposited on 12th May 1998 under
25 the terms of the Budapest Treaty for the International Recognition of Microorganism Deposits for Patent Purposes of 28th April 1977 at the Coleccion Espanola de Cultivos Tipo in plasmid pCLON35 under deposit number CECT5043. BamHI and XhoI, can be used to excise the insert cDNA from this. For in vitro transcription-translation, the full-length TmE2F cDNA was cloned into pBluescriptSK+ using these enzymes.

It will be understood that nucleic acids of the invention may be double stranded DNAs or single stranded DNA of the cDNA or a sequence complementary thereto, eg. such as will have use as a probe.

Preferred nucleic acids are characterised in that they encode for a plant E2F or
5 a functional variant thereof including SEQ ID No 3 or a sequence complementary thereto. Further preferred nucleic acids comprise DNA or RNA of SEQ ID No 5, whether double or single stranded, sense or a sequence complementary thereto. Preferred nucleic acids comprise a cDNA., for example comprising SEQ ID No 3 or 5. Such nucleic acids are optionally provided together with promoter, enhancer or stop
10 sequences with no other gene coding regions. .

The DNA or RNA of the invention may have a sequence containing degenerate substitutions in the nucleotides of the codons in the sequences encoding for E2F proteins or peptides of the invention. In RNA U's replace the T's of DNA. Preferred *per se* DNAs or RNAs are capable of hybridising with the polynucleotides
15 encoding for peptides or proteins of the invention in conditions of low stringency, being preferably also capable of such hybridisation in conditions of high stringency.

The terms "conditions of low stringency" and "conditions of high stringency" are of course understood fully by those skilled in the art, but are conveniently exemplified in US 5202257, columns 9 and 10 and in WO 98/40483 on page 3; both
20 of which are incorporated herein by reference. Thus, generally, the most preferred nucleic acids of the invention will hybridise at the most stringent conditions described in these patents while other embodiments will hybridise at the milder stringency or low stringency conditions.

Where modifications are made they should lead to the expression of a protein
25 with different amino acids in the same class as the corresponding amino acids to these E2F peptide or protein sequences; that is to say, they are conservative substitutions. Such substitutions are known to those skilled in the art see, for example, US 5380712 which is incorporated herein by reference, and are considered only when the protein is active as an E2F peptide or protein as discussed above.

30 The expression 'conservatively substituted' as used with respect to amino acids relates to the substitution of a given amino acid by an amino acid having

physicochemical characteristics in the same class. Thus where an amino acid has a hydrophobic characterising group, a conservative substitution replaces it by another amino acid also having a hydrophobic characterising group; other such classes are those where the characterising group is hydrophilic, cationic, anionic or contains a thiol or thioether. Such substitutions are only contemplated where the resultant protein has activity as an E2F peptide or protein as discussed with respect to DNA and Rb binding.

Nucleic acids of the invention may be degeneratively substituted with respect to that exemplified herein in the sequence listing. The expression 'degeneratively substituted' refers to substitutions of nucleotides by those which result in codons encoding for the same amino acid; such degenerative substitutions being advantageous where the cell or vector expressing the protein is of such different type to the DNA source organism cell that it has different codon preferences for transcription/translation to that of the cDNA source cell. Such degenerative substitutions will thus be host specific.

DNA or RNA provided from a plant or the deposit referred to above may be altered by mutagenic means such as the use of mutagenic polymerase chain reaction primers. Methods of producing the proteins or peptides of the invention characterised in that they comprise use of the DNA or RNA of the invention to express them from cells are also provided in this aspect.

For the purpose of screening for plant E2Fs, a process which has heretofore been hampered due to human E2F dissimilarity to plant E2F, nucleic acid probes or primers comprising a double or single stranded DNA of sequence corresponding to 10 or more contiguous nucleotides taken from the sequence SEQ ID No 5 are provided, with the proviso that they are not selected from those just encoding for the amino acid sequence that is relatively highly conserved with human E2F, ie.

Gln Lys Arg Arg Ile Tyr Asp Ile Thr Asn Val Leu Glu Gly Ile Xaa Leu Ile Glu
Lys Xaa Xaa Lys Asn Xaa Ile Arg Trp

Such probes and primers may be used in Northern and Southern blotting and in PCR, including RT-PCR, and LCR.

Oligonucleotides for use as probes conveniently comprise at least 18 contiguous bases of the sequences of the invention, preferably being of 30 to 100 bases long, but may be of any length up to the complete sequence or even longer. For use as PCR or LCR primers the oligonucleotide preferably is of 10 to 20 bases long but may be longer. Primers should be single stranded but probes may be also be
5 double stranded ie. including complementary sequences.

For the purpose of downregulating native plant E2F expression there is also provided antisense DNA to any of the nucleic acids of the invention described above. This technique is well known in the art but is generally illustrated by US 5356799 and
10 US 5107065 by way of example, each of which is incorporated herein by reference.

A fourth aspect of the invention provides a nucleic acid vector or construct comprising a nucleic acid of the present invention or comprising antisense nucleic acid thereto. Suitable vectors or constructs for introducing the peptides or proteins of the invention into plants will occur to those skilled in the art of plant molecular
15 biology, but are conveniently those discussed below with respect to methods for producing transgenic plants.

A fifth aspect of the present invention provides a plant cell comprising recombinant nucleic acid, preferably recombinant DNA, of the third aspect of the invention. Nucleic acids of the invention are particularly provided in the form of such
20 nucleic acid vectors or DNA construct comprising that nucleic acid or antisense nucleic acid sequence thereto.

A sixth aspect of the present invention provides a plant cell comprising antisense nucleic acid thereto capable of downregulating expression of native plant E2F.

25 A seventh aspect of the present invention comprises a transgenic plant or part thereof comprising recombinant nucleic acid, a vector or DNA construct as described above.

It will be realised that a most effective method of delivering proteins and peptides of the invention to plant cells is by expressing nucleic acid encoding *them in situ*. Such method is conventionally carried out by incorporating oligonucleotides or
30 polynucleotides, having sequences encoding the peptide or protein, into the plant cell

DNA. Such nucleotides can also be used to downregulate native E2F expression by gene silencing coexpression or through antisense strategy. By use of mutagenesis techniques, eg. such as SDM, the nucleotides of the invention may be designed and produced to encode proteins and peptides which are functional variants or otherwise
5 overactivated or inactivated, eg. with respect to binding, of the invention

Preferred plants of the seventh aspect may comprise the nucleic acid of the invention in a construct in functional association with promoter, activating or otherwise regulating sequences. Preferred promoters may be tissue specific such that the resultant expression of peptide, and thus its effects, are localised to a desired
10 tissue. Promoters with a degree of tissue specificity will be known to those skilled in the art of plant molecular biology. Some of these are discussed below.

Methods of producing vectors and constructs capable of being used in the present invention will occur to those skilled in the art in the light of conventional molecular biology techniques. DNA, RNA and vector containing or encoding for
15 these may be introduced into target cells in known fashion in the field of plant cell transformation. Particularly preferred is the method of introducing the DNA or RNA into pollen cells using techniques such as electroporation or gene gun technology.

It may be preferred to express the DNA or RNA of the invention throughout the plant, but in the event that tissue specific effect is to be exploited then it will be
20 understood by those skilled in the art that tissue specific promoters, enhancers or other activators should be incorporated into the transgenic cells employed in operative relation with the DNA.

It will be realised by those skilled in the art that suitable promoters may be active ectopically, continuously or may be inducible. It will be appreciated by those
25 skilled in the art that inducible or tissue specific ie promoters will have advantage in so far as they are capable of providing alteration of the aforesaid E2F peptide or protein activity only when or where required, eg. at a particular stage of cell development or in a tissue such as leaves, roots, fruit or seeds or subparts thereof, eg. endosperm, that may be the subject of desired increase or decrease in size or even
30 deletion.

No particular limitation on the type of promoter to be employed is envisioned, although a reasonable amount of experimental trial may be expected to be undertaken to produce good results. Examples of tissue specific and inducible promoters can be found in the following patent literature: US 5086169 (pollen specific), US 5459252 and US 5633363 (root specific), US 5097025 ((i)seed, (ii)mature plant), US 5589610 (stamen), US 5428146 (wound), US 5391725 ((i)chloroplast, (ii) cytosol), US 4886753 (root nodule), US 4710461 (pollen), US 5670349 (pathogen), US 5646333 (epidermis), US 5110732 ((i) root , (ii) radical), US 5859328 (pistil), US 5187267 (heat shock), US 5618988 (storage organ), US 5401836 and US 5792925 (root), US 4943674 (fruit), US 5689044 and US 5654414 (chemical), US 5495007 (phloem), US 5589583 (meristem), US 5824857 (vasculature), each of which is incorporated herein by reference. Constitutive promoters will be well known to those skilled in the art and are discussed in the documents above and referred to below but for example include CaMv35S and alfalfa (MsH3g1) (see WO 97/20058 incorporated herein by reference).

Numerous specific examples of methods used to produce transgenic plants by the insertion of cDNA in conjunction with suitable regulatory sequences will be known to those skilled in the art. Plant transformation vectors have been described by Denecke et al (1992) EMBO J. 11, 2345-2355 and their further use to produce transgenic plants producing trehalose described in US Patent Application Serial No. 08/290,301. EP 0339009 B1 and US 5250515 describe strategies for inserting heterologous genes into plants (see columns 8 to 26 of US 5250515). Electroporation of pollen to produce both transgenic monocotyledonous and dicotyledonous plants is described in US 5629183, US 7530485 and US 7350356. Further details may be found in reference works such as Recombinant Gene Expression Protocols. (1997) Edit Rocky S. Tuan. Humana Press. ISBN 0-89603-333-3; 0-89603-480-1. All of these documents are incorporated herein by reference. It will be realised that no particular limitation on the type of transgenic plant to be provided is envisaged; all classes of plant, monocot or dicot, may be produced in transgenic form incorporating the nucleic acid of the invention such that E2F activity in the plant is altered, constitutively or ectopically.

In an eighth aspect of the present invention the present inventors have provided antibodies capable of specifically binding with plant E2F factor peptides or proteins of the first aspect of the present invention, thus enabling the identification and isolation of further peptides and proteins of the invention and nucleic acid sequences encoding therefor, eg. using techniques such as Western blotting.

The present invention will now be illustrated further by reference to the following non-limiting Examples. Further embodiments falling within the scope of the claims attached hereto will occur to those skilled in the art in the light of these.

FIGURES.

Fig. 1. DNA sequence of the wheat cDNA encoding E2F protein and deduced amino acid sequence.

Fig. 2. Northern analysis to identify mRNA encoding wheat E2F.

Fig. 3. Amino acid alignment of wheat E2F with human and Drosophila E2F proteins.

Fig. 4. Interaction between plant retinoblastoma protein (ZmRb1) and plant E2F protein by yeast two-hybrid analysis.

Fig. 5. Domain organization of human E2F-1 and wheat E2F proteins.

EXAMPLES

EXAMPLE 1: Isolation of plant E2F cDNA clone

To identify proteins which interact with plant Rb, we carried out a yeast two-hybrid screening of a wheat cDNA library made from proliferating wheat cells growing in suspension culture. A large number of positive interactors were recovered, which allowed yeast co-transformants to grow under highly stringent conditions (20-30 mM 3AT) and to yield a positive β -gal signal. DNA sequencing analysis revealed that two of the strong interactors contained cDNA inserts of ~1.1 kb and had identical DNA sequences. When this DNA sequence was used as a query in a BLAST search, several members of the E2F family were retrieved. In particular, the deduced amino acid sequence of the isolated cDNA clone showed a significant homology with the heterodimerization domain of human E2F-5. The cDNA as well as an oligonucleotide derived from its 5' end were used to screen a wheat cDNA

library by colony hybridization. Four positive clones, containing inserts of ~2.0 kb, were recovered. The sequence of the longest cDNA insert, shown in Figure 1, contains a single ORF of 1371 bp, with the potential to encode a protein of 458 amino acids. This ORF is flanked by 170 bp and 439 bp of 5' and 3' untranslated regions, respectively.

The plant Rb-interacting cDNA clone encodes a plant homologue of animal E2F. Northern analysis indicated that a message, ~2.0 kb in length, with the capacity to encode the entire TmE2F ORF, is present in RNA prepared from shoots and leaves, where most of the cells do not proliferate, as well as from root meristems and proliferating suspension cultured cells (Fig. 2). With the study presented here, we can not fully rule out the possibility that other, more distantly E2F-related genes, may exist. So far, Southern analysis strongly suggests that wheat E2F is the product of a single copy gene. In vitro transcription-translation reactions programmed with a plasmid containing the entire TmE2F cDNA insert yielded a major product with a mobility corresponding to ~58-60 kDa apparent molecular mass (Fig. 2), slightly larger than predicted from the deduced amino acid sequence.

The idea that the TmE2F cDNA clone encodes a plant E2F protein homologous to the animal counterparts is reinforced by analysis of the amino acid homology and domain organization of plant E2F. Based on a pairwise distance analysis, obtained with the CLUSTAL algorithm, plant E2F exhibits an overall ~24.0-27.5 % amino acid similarity with the subset formed by human E2F-1 (Helin et al., 1992; Kaelin et al., 1992; Shan et al., 1992), E2F-2 (Ivey-Hoyle et al., 1993; Lees et al., 1993) and E2F-3 (Lees et al., 1993), a slightly larger similarity (~25.0-29.8%) with E2F-4 (Beijersbergen et al., 1994; Ginsberg et al., 1994; Sardet et al., 1995) and E2F-5 (Sardet et al., 1995), and a much lower similarity (18.8%) with *Drosophila* E2F (Dymlacht et al., 1994; Ohtani et al., 1994).

Amino acid alignment of plant and animal E2F proteins revealed a similar domain organization and some specific characteristics of plant E2F. The most conserved domain appears to be the DNA binding domain which is highly homologous among all members (Fig. 3). This domain includes a stretch of 15 amino acids (residues 182-196) fully conserved, which corresponds to one of the putative α

helices of the conserved bHLH domain (Cress et al., 1993). A significant degree of conservation between plant and animal E2F proteins was also found within the homo- and heterodimerization domains, including the characteristic leucine zipper motif (residues 219-240). However, based on this homology analysis, a typical cyclin A box characteristic of some human members, was not apparent in plant E2F. Similarly, the nuclear localization signal (NLS), typical of human E2F-1, E2F-2 and E2F-3 is not found, in the same location, in plant E2F. However, a short amino stretch (residues 74-76), located in a more N-terminal position than in human E2F, which may act as NLS is present plant E2F (Fig. 3).

An interesting characteristic of plant E2F is that the homology within the C-terminal third of the protein, containing the transactivation and Rb-binding domains in human E2F members, is very reduced at the level of primary sequence. In particular, the sequence of the C-terminal 18 amino acids which confer Rb-binding ability to human E2Fs, is not present in plant E2F, although its C-terminal residues are required for Rb binding (see below). However, a manual adjustment of the alignment output allows the identification of a 16 amino acid motif in plant E2F (YX₆DX₄DMWE; positions 407-422) which may be homologous to the Rb-binding motif of animal E2Fs, which is fully conserved in members of all animal species. Interestingly, a similar spacing between critical amino acids as well as a conservation of the acidic and hydrophobic nature of some critical residues, strongly supports our proposal that it may represent the minimal Rb-binding motif of plant E2F.

EXAMPLE 2: Protein domains required for E2F/ZmRb1 interaction

To investigate the amino acid requirements for the interaction between plant E2F and Rb, we carried out a yeast two-hybrid analysis using several truncated proteins. Human Rb and related proteins bind to E2F family members using their A7B pocket domain (Lees et al., 1993). To establish the protein domain required in plant Rb to interact with plant E2F, yeast cells were cotransformed with plasmids expressing the Gal4AD-E2F fusion protein and plasmids expressing the Gal4BD alone or fused to several truncated versions of plant Rb. Cells were grown on plates with and without histidine supplemented with 3-AT, as indicated in Fig. 4. Growth

on plates lacking histidine was entirely dependent on the presence and interaction of both plant Rb and E2F protein. Deletion of the 125 C-terminal residues of plant Rb (ZmRb- Δ C2) did not markedly reduce protein interaction, as it was also the case with a truncated Rb protein containing the A/B pocket and the C-terminal domain (ZmRb- Δ N). The A/B pocket alone (ZmRb1- Δ N Δ C2) was able to support interaction, although with a slightly reduced efficiency (Fig. 4). These growth characteristics of the yeast cotransformants correlated well with expression of β -galactosidase activity.

A similar study was carried to determine the region in plant E2F involved in Rb binding. In human E2F, pocket proteins bind to the C-terminal residues (reviewed in Slansky and Farnham, 1996). Yeast cotransformants expressing a truncated plant E2F (236-458) were able to grow in the absence of histidine (Fig. 4). However, elimination of the C-terminal residues (TmE2F 236-373) did not allow growth in the absence of histidine (Fig. 4). This indicates that C-terminal domain of plant E2F contains the Rb-binding motif. Moreover, these C-terminal residues involved in plant E2F-Rb interaction contains the 16 amino acid motif identified in this study (see alignment in Fig. 3). Altogether, these studies lead us to conclude that plant E2F represents a novel member of the E2F family of transcription factors in which several degrees of amino acid conservation can be recognized in the different protein domains.

EXAMPLE 3: Plant E2F: domain organization and properties

A comparison of the domain organization of plant and human E2F proteins is shown in Fig. 5.

The DNA binding domain.

Based on mutational analysis, human E2F-1 was originally described as a basic helix-loop-helix (bHLH) protein (Cress et al., 1993). The DNA binding domain of plant E2F (residues 146-209) is the most conserved region of the protein not only with mammalian E2F members but also with *Drosophila* E2F. Based on this high degree of conservation, one prediction is that plant E2F should bind to a DNA sequence very similar to the consensus human E2F-binding site

(TTT(C/G)(C/G)CG(C/G); reviewed in Cobrinik, 1996). Among the plant promoters which have been cloned and sequenced, E2F-consensus binding sequences have been found in the ribonucleotide reductase genes of *Nicotiana tabacum* (C. Gigot, personal communication).

5

The Rb binding motif

One striking feature of plant E2F is the low amino acid similarity of in the C-terminal region, which contains its Rb-binding motif, in relation to the high homology of other domains, e. g. the DNA-binding domain, among all animal E2Fs. It has been found that amino acids 409-426 in the C-terminal domain of human E2F-1, containing a relatively high proportion of acidic residues, are sufficient for binding to Rb and that point mutations within this short region drastically modify the ability of human E2F-1 to associate with Rb (Cress et al., 1993, Helin et al., 1993). Among them, we can find the 16 amino acid motif YX₇EX₃DLFD (positions 411 to 426 in human E2F-1), absolutely conserved in all animal E2Fs (see also Fig. 3), which has been shown to be critical for E2F-1 binding to Rb in human cells (Shan et al., 1996). Plant E2F contains a 16 amino acid motif (YX₆DX₄DMWE; positions 407-422). Interestingly, a similar spacing between critical amino acids as well as a conservation of acidic and hydrophobic residues, strongly supporting our proposal that it directs binding to plant Rb.

20

The putative NLS of plant E2F.

It has been recently shown that transcriptional activity of human E2F appears to be finely regulated by changes in the subcellular localization (Verona et al., 1997). In fact, E2F-1, -2 and -3 contains a short stretch of amino acids, absent in E2F-4, which act as a nuclear localization signal (NLS) and is related to that of c-myc (Dang and Lee, 1988). Plant E2F does not contain such a consensus sequence. Therefore, we can speculate that plant E2F is translocated to the nucleus by other partner proteins. Alternatively, a different NLS may be present in plant E2F. In fact, the region of plant E2F encompassing residues 69 to 81 (RTRQLKRKATREE) may behave as a NLS. It is important to mention that maize Rb has been shown to have largely a

30

nuclear localization (Ach et al., 1997). Since a clear NLS is not apparent in maize Rb, it may occur that the E2F NLS targets Rb into the nucleus of the plant cell in the Rb/E2F complex. Exclusion of E2F from this complex may be a regulatory way to exclude Rb from the nucleus. It is now proposed that this is a way to avoid Rb repression of genes.

MATERIAL AND METHODS

DNA manipulations and plasmid constructions.

Standard DNA manipulation techniques were applied as described in
10 Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

DNA sequencing was carried out using an Applied Biosystem 373A device. Oligonucleotides were from Isogen Bioscience BV (Maarsen, The Netherlands). Plasmid pGBT-ZmRb1 was constructed by cloning the ZmRb1 cDNA (15) in frame to the Gal4^{BD} of pGBT8, pGBT-ZmRb1ΔC2(1-558) by deleting a MscI-XhoI fragment of pGBT-ZmRb1 and pGBT-ZmRb1ΔNΔC2(69-558) by deleting a MscI-XhoI fragment of pGBT-ZmRb1ΔN. Plasmid pGBT-ZmRb1ΔN(69-683) contains a N-terminal deletion of ZmRb1. Plasmid pGADTmE2F(236-458) is a partial clone isolated in the screening and pGADTmE2F(236-373) was made by deleting a SspI-XhoI fragment. For *in vitro* transcription-translation, the full-length TmE2F cDNA
20 was cloned into pBluescriptSK+. Plasmids pGADE2F-1 and pGADE2F-5, containing human E2F-1 and E2F-5, respectively, were provided by N. LaThangue and S. de la Luna, and plasmids p130Rb2 (20) and pGT-RB (21) by M. Serrano.

25 Construction of the yeast two-hybrid cDNA library from wheat cultured cells

Five micrograms of poly(A)⁺ mRNA isolated from wheat suspension cultured cells were used as a substrate for cDNA synthesis using a cDNA synthesis kit (Stratagene), according to the manufacturer's instructions. The resulting double-stranded DNA, containing EcoRI and XhoI ends, had an average size of 1.3 Kb. A
30 sample (500 ng) of this cDNA was ligated to 750 ng of the EcoRI/XhoI-digested pGAD-GH vector (Clontech) for 48 hr at 8°C. Following ligation, the library was

dialyzed against distilled water and electroporated into *E. coli* DH10B (Gibco). Total library DNA was obtained by plating primary transformants on fifty 150-mm LB plates plus ampicillin. Colonies were scrapped off into LB (+Amp) medium, and plasmid DNA was prepared as described in Sambrook.

5

Yeast two-hybrid screening

The yeast strain HF7c (MATa *ura3-52 his3-200 ade2-101 lys2-801 trp1-901 leu2-3,112 gal4-542 gal80-538 LYS2::GAL1UAS-GAL1TATA-HIS3 URA3::GAL4* 17mers(x3)-*CYC1TATA-LacZ*; Feilotter *et al* 1994, which contains the two reporter genes *LacZ* and *HIS3*, was used in the two-hybrid screening. Yeasts were first transformed, with pGBTZmRb1, a plasmid containing the maize Rb protein (Xie *et al.*, 1996) fused to the Gal4 DNA-binding domain (BD; TRP1 marker) in the pGBT8 vector. Then, they were transformed with the pGAD-GH (AD; LEU2 marker) wheat cDNA library. The transformation mixture was plated on yeast drop-out selection media lacking tryptophan, leucine and histidine and supplemented with 5 mM and 10 mM 3-amino-1,2,4, triazole (3-AT) to reduce the appearance of false positive growing colonies. Transformants were routinely recovered during a 3 to 8 days period and were checked for growth in the presence of up to 20 mM 3-AT. To corroborate the interaction between the two fusion proteins, β -galactosidase activity was assayed by a replica filter assay as described. Plasmid DNA was recovered from positive colonies by transforming into *E. coli* MH4, since this strain is *leuB*⁻, and its defect can be complemented by the *LEU2* gene present in the pGAD-GH plasmid.

Purification of GST fusion proteins and *in vitro* transcription and translation.

E. coli BL21(DE3) transformants were grown to an OD₆₀₀ of 0.6-0.9 and induced with 1 mM IPTG. GST fusion proteins were purified using glutathione-Sepharose beads (Pharmacia). ³⁵S-methionine labeled Tme2F protein was obtained by using the TNT kit (Promega).

Wheat cell cultures.

The *Triticum monococcum* suspension culture (P. M. Mullineaux; John Innes Centre, UK), was maintained as described (13). Cells were synchronized with 10 mM hydroxyurea (HU) for 48 hours.

5

Northern and Southern analysis.

Ten micrograms of total wheat cell RNA were denatured, fractionated in a 1.2% agarose gel plus 2.2 M formaldehyde, and transferred to a Zeta-Probe membrane (Bio-Rad). The TmE2F (nt 935-1635) and wheat histone H4 (Xie and
10 Gutierrez, unpublished) probes were labeled by random priming with α -³²P-dCTP, and mixed for hybridization. Ten μ g of genomic wheat DNA was digested with the indicated enzymes, fractionated in 0.8% agarose gels, transferred to BioDyne (Amersham) membranes and probed as described in Sambrook et al.

15 EXAMPLE 4: Production of antibodies specific for binding to plant E2F protein.

Polyclonal antibodies capable of specifically binding plant E2F protein were provided by producing a GST fusion with the 236-458 C-terminal fragment in Bluescript as described above. This was over-expressed in *E.coli* and purified on a Glutathione bead column. Rats were injected using standard immunisation protocols
20 on day 1 and day 14 and serum derived from these used as polyclonal reagent. This serum was capable of use at 1/1000 dilution for Western Blotting purposes. (see standard procedures in Manual of Antibody Preparation. Coldspring Harbor Press).

REFERENCES- incorporated herein by reference.

1. Nevins, J.R. (1992) *Science* 258, 424-429.
- 25 2. Helin, K. (1998) *Curr. Opin. Genet. Dev.* 8, 28-35.
3. Weinberg, R.A. (1995) *Cell* 81, 323-330.
4. Brehm, A., Miska, E.A., McCance, D.J., Reid, J.L., Bannister, A.J. & Kouzarides, T. (1998) *Nature* 391, 597-601.

5. Magnaghi-Jaulin, L., Groisman, R., Naguibneva, I., Robin, P., Lorain, S., Le Villain, J.P., Troalen, F., Trouche, D. & Harel-Bellan, A. (1998) *Nature* **391**, 601-605.
6. Staiger, C. & Doonan, J. (1993) *Curr. Opin. Cell Biol.* **5**, 226-231.
- 5 7. Doonan, J. & Fobert, P. (1997) *Curr. Opin. Cell Biol.* **9**, 824-830.
8. Philipps, G., Clement, B. & Gigot, C. (1995) *FEBS Lett.* **358**, 67-70.
9. Shen, W.H. & Gigot, C. (1997) *Plant Mol. Biol.* **33**, 367-379.
10. Fuerst, R.A., Soni, R., Murray, J.A.H. & Lindsey, K. (1996) *Plant Physiol.* **112**, 1023-1033.
- 10 11. Soni, R., Carmichel, J.P., Shah, Z.H. & Murray, J.A.H. (1995) *Plant Cell* **7**, 85-103.
12. Dahl, M., Meskiene, I., Bögre, L., Ha, D.T.C., Swoboda, I., Hubmann, R., Hirt, H. & Heberle-Bors, E. (1995) *Plant Cell* **7**, 1847-1857.
13. Xie, Q., Suarez-Lopez, P. & Gutierrez, C. (1995) *EMBO J.* **14**, 4073-4082.
- 15 14. Grafi, G., Burnett, R.J., Helentjaris, T., Larkins, B.A., DeCaprio, J.A., Sellers, W.R. & Kaelin, W.G., Jr. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 8962-8967.
15. Xie, Q., Sanz-Burgos, A.P., Hannon, G.J. & Gutierrez, C. (1996) *EMBO J.* **15**, 4900-4908.
16. Ach, R.A., Durfee, T., Miller, A.B., Zambryski, P.C., Hanley-Bowdoin, L. & Gruissem, W. (1997) *Mol. Cell. Biol.* **17**, 5077-5086.
- 20 17. Huntley, R., Healy, S., Freeman, D., Lavender, P., de Jager, S., Greenwood, J., Makkerh, J., Walker, E., Jackman, M., Xie, Q., Bannister, A.J., Kouzarides, T., Gutierrez, C., Doonan, J.H. & Murray J.A.H. (1998) *Plant Mol. Biol.* **37**, 155-69.
18. Koguni, S., Suzuka, I., Ohashi, Y., Murakami, T & Arai, Y. (1991) *Nucleic Acids Res.* **19**, 1571-1576.
- 25 19. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor,

20. Hannon, G.J., Demetrick, D. & Beach, D. (1993) *Genes & Dev.* 7, 2378-2391.
21. Ewen, M.E., Sluss, H.K., Sherr, C.J., Matsushime, H., Kato, J. and Livingston, D.M. (1993) *Cell* 73, 487-497.
22. Fields, S. & Song, O. (1989) *Nature* 340, 245-246.
- 5 23. Breeden, L. & Nasmyth, K. (1985) *Cold Spring Harbor Symp. Quant. Biol.* 50, 643-650.
24. Helin, K., Lees, J.A., Vidal, M., Dyson, N., Harlow, E. & Fattaey, A. (1992) *Cell* 70, 337-350.
25. Kaelin, W.G., Jr., Krek, W., Sellers, W.R., DeCaprio, J.A., Ajchenbaum, F.,
10 Fuchs, C.S., Chittenden, T., Li, Y., Farnham, P.J., Blunar, M.A., Livingston, D.M. & Flemington, E.K. (1992) *Cell* 70, 351-364.
26. Shan, B., Zhu, X., Chen, P.-L., Durfee, T., Yang, Y., Sharp, D. & Lee, W.-H. (1992) *Mol. Cell. Biol.* 12, 5620-5631.
27. Ivey-Hoyle, M., Coroy, R., Huang, H.E., Goodhart, P.J., Oliff, A. and
15 Heimbrook, D.C. (1993) *Mol. Cell. Biol.* 13, 7802-7812.
28. Lees, J.A., Saito, M., Vidal, M., Valentine, M., Look, T., Harlow, E., Dyson, N. & Helin, K. (1993) *Mol. Cell. Biol.* 13, 7813-7825.
29. Beijersbergen, R.L., Kerkhoven, R.M., Zhu, L., Carlée, L., Voorhoeve, P.M. & Bernards, R. (1994) *Genes & Dev.* 8, 2680-2690.
- 20 30. Ginsberg, D., Vairo, G., Chittenden, T., Xiao, Z-X, Xu, G., Wydner, K.L., DeCaprio, J.A., Lawrence, J.B. & Livingston, D.M. (1994) *Genes & Dev.* 8, 2665-2679.
31. Sardet, C., Vidal, M., Cobrinik, D., Geng, Y., Onufryk, C., Chen, A. & Weinberg, R.A. (1995) *Proc. Natl. Acad. Sci. USA* 92, 2403-2407.
- 25 32. Dynlacht, B.D., Brook, A., Dembski, M., Yenush, L. & Dyson, N. (1994) *Proc. Natl. Acad. Sci. USA* 91, 6359-6363.
33. Ohtani, K. & Nevins, J.R. (1994) *Mol. Cell. Biol.* 14, 1603-1612.

34. Cress, W.D., Johnson, D.G. & Nevins, J.R. (1993) *Mol. Cell. Biol.* 13, 6314-6325.
35. Cobrinik, D. (1996) *Curr. Topics Microbiol. Immunol.* 208, 31-61.
36. de la Luna, S., Burden, M.J., Lee, C.W. & LaThangue, N.B. (1996) *J. Cell Sci.* 109, 2443-2452.
37. Verona, R., Moberg, K., Estes, S., Starz, M., Vernon, J.P. & Lees, J.A. (1997) *Mol. Cell. Biol.* 17, 7268-7282.
38. Slansky, J.E. & Farnham, P.J. (1996) *Curr. Topics Microbiol. Immunol.* 208, 1-30.
39. Helin, K., Harlow, E. and Fattaey, A. (1993) *Mol. Cell. Biol.* 13, 6501-6508.
40. Shan, B., Durfee, T. & Lee, W.H. (1996) *Proc. Natl. Acad. Sci. USA* 93, 679-684.
41. Lee, J.-O., Russo, A.A. & Pavletich, N.P. (1998) *Nature* 391, 859-865.
42. Lam, E.W.-F & LaThangue, N.B. (1994) *Curr. Opin. Cell Biol.* 6, 859-866.
43. Magyar, Z., Meszaros, T., Miskolczi, P., Deak, M., Feher, A., Brown, S., Kondorosi, E., Athanasiadis, A., Pongor, S., Bilgin, M., Bako, L., Koncz, C. & Dudits, D. (1997) *Plant Cell* 9, 223-235.
44. Wang, H., Fowke, L.C. & Crosby, W.L. (1997) *Nature* 386, 451-452.
45. Hirt, H., Pay, A., Bogre, L., Meskienc, I. & Heberle-Bors, E. (1993) *Plant J.* 4, 61-69.
46. Fobert, P.R., Gaudin, V., Lunness, P., Coen, E.S. & Doonan, J. (1996) *Plant Cell* 8, 313-320.
47. Brook, A., Xie, J. E., Du, W. & Dyson, N. (1996) *EMBO J.* 15, 3676-3683.
48. Knudsen, E.S. & Wang, J.Y. (1997) *Mol. Cell. Biol.* 17, 5771-5783.
49. Loidl, A. & Loidl, P. (1996) *Crit. Rev. Oncogen.* 7, 49-64.

CLAIMS

1. A method of controlling plant growth and/or cellular DNA replication and/or cell cycle progression, differentiation and development comprising increasing or
5 decreasing E2F activity in a plant cell through expression of a recombinant plant E2F peptide or protein in that cell.
2. A method as claimed in Claim 1 characterised in that the E2F activity comprises one or both of (i) the ability to bind plant Retinoblastoma protein and (ii)
10 the ability to bind to E2F transcription factor binding sites in plant DNA.
3. A method as claimed in Claim 1 characterised in that it comprises altering the plant E2F protein level, subcellular localisation, the E2F DNA-binding activity, the E2F protein-protein binding activity, the E2F transactivation properties, and/or the
15 E2F-Rb-binding activity.
4. A method as claimed in any one of Claims 1 to 3 characterised in that plant E2F
is modified alone and/or in combination with a modification of the levels or activity
20 of plant Rb.
5. A method as claimed in any one of the preceding claims characterised in that it alters cell shape.
- 25 6. A method as claimed in any one of the preceding claims characterised in that it alters cell proliferation characteristics such as to increase plant cell or plant organ size.
7. A method as claimed in any one of the preceding claims characterised in that it
30 increases or decreases expression of other proteins.

8. An isolated, enriched, cell free and/or recombinantly produced protein or peptide characterised in that it has one or both plant E2F activities selected from (i) the ability to bind plant Retinoblastoma protein and (ii) the ability to bind to E2F transcription factor binding sites in plant DNA

5 wherein the protein or peptide comprises one or both amino acid sequences selected from the following domains of SEQ ID No 6:

(a) Tyr Xaa Xaa Xaa Xaa Xaa Xaa Asp Xaa Xaa Xaa Xaa Asp Met Trp Glu

10 (b) Gln Lys Arg Arg Ile Tyr Asp Ile Thr Asn Val Leu Glu Gly Ile Xaa Leu Ile Glu
Lys Xaa Xaa Lys Asn Xaa Ile Arg Trp

provided that where the peptide or protein comprises only domain (b) it comprises at least 50% of the contiguous sequence of amino acids 1-406 of SEQ ID No 6.

15

9. A peptide or protein as claimed in Claim 8 further characterised in that it comprises a sequence

(c) Leu Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Glu Xaa Xaa Xaa Leu Xaa Xaa Xaa
20 Glu Xaa Xaa Leu Asp

10. A peptide or protein as claimed in Claim 8 or Claim 9 characterised in that it is of 16 to 100 aminoacids.

25 11. A peptide or protein as claimed in any one of Claims 8 to 10 characterised in that it comprises an amino acid sequence of SEQ ID No 2 or a functional variant thereof.

30 12. A peptide or protein as claimed in any one of Claims 8 to 11 characterised in that it is of SEQ ID No 6 but modified such that the amino acid sequence SEQ ID No 2 is mutated such that its ability to bind Rb protein is reduced from that of the native

sequence of SEQ ID No 2 or abolished completely therefrom, whereby the peptide is capable of acting as an E2F protein without being restricted by Rb binding.

13. A peptide or protein as claimed in any preceding claim characterised in that it
5 further comprises a sequence of SEQ ID No 7, that being of sequence

Arg Thr Gln Leu Lys Arg Lys Ala Thr Arg Glu Glu

or a functional variant thereof having functional activity as a nuclear localisation signal (NLS).

- 10 14. A peptide as claimed in any one of the preceding claims in which the NLS of SEQ ID No 7 is modified such that the peptide does not localise in the nucleus.

- 15 15. A peptide as claimed in any one of the preceding Claims 8 to 14 characterised in that it comprises a DNA binding domain being of sequence of amino acid residues 146-206 of the plant E2F of SEQ ID No 6 or a functional equivalent thereof.

16. A peptide as claimed in any of Claims 8 to 15 characterised in that it binds to an E2F DNA binding site of sequence TTT(C/G)(C/G)(C/G)(C/G)(C/G).

- 20 17. An isolated, enriched, cell free and/or recombinant nucleic acid comprising a sequence encoding for expression of a peptide or protein as described in any one of Claims 8 to 16 or a sequence complementary thereto.

- 25 18. A nucleic acid as claimed in claim 16 characterised in that it encodes for a plant E2F peptide or protein or a functional variant thereof including SEQ ID No 3 or a sequence complementary thereto.

19. A nucleic acid as claimed in Claims 17 or 18 comprising DNA or RNA of SEQ ID No 5 or a sequence complementary thereto.

30

20. A nucleic acid as claimed in any one of claims 17 to 19 characterised in that it comprises a cDNA.
21. A nucleic acid as claimed in claim 19 characterised in that it comprises a SEQ
5 ID No 3 or 5 optionally together with promoter, enhancer or stop sequences but no other gene coding regions.
22. A nucleic acid probe or primer comprising a double or single stranded DNA of sequence corresponding to 10 or more contiguous nucleotides taken from the
10 sequence SEQ ID No 5 provided that they are not selected from those encoding for the amino acid sequence
Gln Lys Arg Arg Ile Tyr Asp Ile Thr Asn Val Leu Glu Gly Ile Xaa Leu Ile Glu Lys
Xaa Xaa Lys Asn Xaa Ile Arg Trp.
- 15 23. A nucleic acid vector or construct comprising a nucleic acid as claimed in any one of claims 16 to 20 or comprising an antisense nucleic acid sequence thereto.
24. A plant cell comprising recombinant nucleic acid as claimed in any one of
20 claims 16 to 20.
25. A plant cell comprising antisense nucleic acid to plant E2F expressing nucleic acid capable of downregulating expression of native plant E2F as claimed in any one of Claims 16 to 20.
- 25 26. A transgenic plant or part thereof comprising a peptide, protein, recombinant nucleic acid, a vector or construct as claimed in any one of the preceding claims.
27. A transgenic plant characterised in that it expresses an E2F protein ectopically, expresses an E2F protein or peptide that inhibits binding of plant Rb protein to native
30 E2F protein or an E2F protein that is resistant to the effects of plant Rb protein.

28. A plant as claimed in claim 25 characterised in that the E2F is of SEQ ID No 6 or is a functional variant thereof.

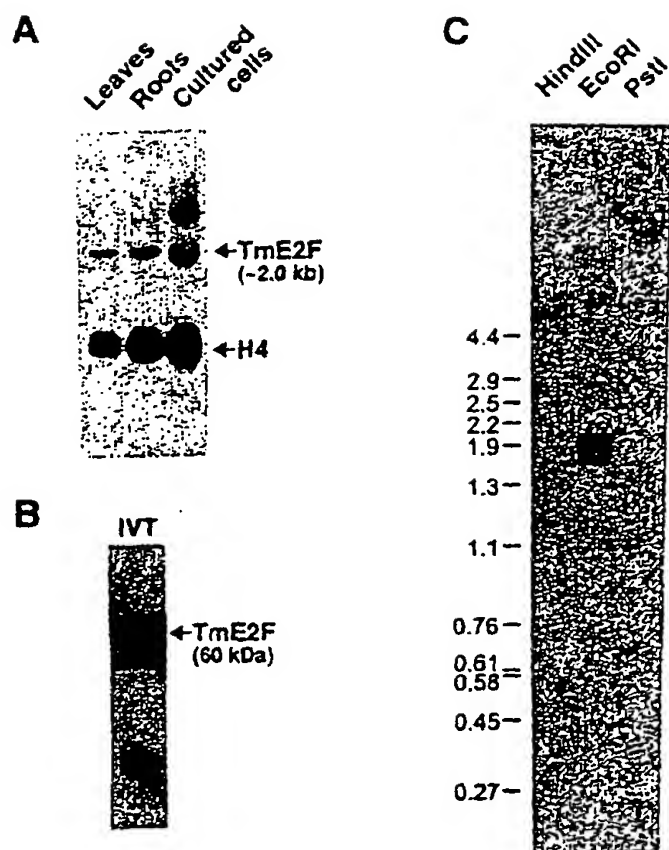
29. A method of producing a plant or plant cell or plant part characterised in that it
5 comprises introducing a nucleic acid as claimed in any one of claims 17 to 23 into a plant cell

30. An antibody characterised in that it binds to a peptide or protein as claimed in any one of Claims 8 to 16.

10

AATTGGGACGAGGCCACCCACCTACCTCCCGCCGCCGCCGCCGCGCCACGGGAACCTATGCTCCGGGAGCCCCCCCCGATGGCCTCTCTGCTTTCTCTTGAAGCCGAAGACGCCCATCGCTC	120
CCGGGAGTCGGGGGTCCCGTAGCGCGCGATCGCAGATCGGGCTTATGCTCGGGGGGGCAGGCGCGCGCGCTGGCAAAAAATATCTGACGATCTCTGCGCCGCCGCCCGGGTGTCTCCACG	240
M S G G R P P A A K I L Q L S L R P P V G F T	25
CCGTCGGGGCTCCCTTCCCTCACCCTCAGCAGTACCAACCGCTTTCATGCGCGGACTACCCCTTCTGCCACTGGCTCCGGCGGCATCGGCTCCGGTGGTGTGGCGCGGATATTGATAG	360
P S R P P F A S P D D Y H R F H A P T T P S A T G S G G I G S G G V V G G D I D E	65
GGGCTTGTTATCCGACGACGCTAAAAAGAAAGCCACACGCGAAGAAAATAATCGCGCTAGCTCAGTGTCTGATTTGTGTACCACTGGAGTTACTGGCAATCCGCTACTCCGCA	480
G L V I R T Q L K R K A T C R E S E N N A A E S S D C M A I V T T G V T G C N P L L T C P	105
GTGTCGGAAGAGTTGTAAGAATTCTAAATCAAGACTAAGAACAATAAAGCTGGCGCTGACACACCTACCGCAAAATGTTGGTCCACCACCTAATCCATCAACTCTCTCTGCTACTTCG	600
V S G X A V K N S K S K T N N N K A G P Q T P T P N V C S P L N P S T P A G T C	145
CGCTATGACAGTTCGTTAGGACTTCTGACAAAGAGTTCATCAATTTGGTGAAGCAAGCTGAGGATGGCATTCTAGATTGCAATATGCTGCAGAAACACTAGAGGTTCAAAAGCGAGCG	720
R Y D S S L G L L T K K F I N L L K Q A E D G I L D L N H A A E T L E V Q E R R	185
ATATATGACATCACAAATGCTCCGAAGGATTTGGTCTTATAGAAAGACACCTTAAACACAGAATTCGTTGAAGGGCTTGATCATTCAGGAGTGAATTAGATATGGCCTTCAGGT	840
I Y D I T T N V L E G I G L I E K T L K N R I R W K G L D D S G V E L D N G L S G	225
TTGACAGAGAAGTTGAAAAATCTTAATTTGCAGGACGAGCCTTAGATGACCGTATTAAGTGATATGCGCGAAAAGCTAAGGGGGTTAACCGAAGATGAGAACATCAAAAGATGGCTCTAT	960
L Q T C T E V E N L N L Q E Q A L D E R I S U M R E K L R G L T E D E N S G C R W L Y	265
GTGACCGAAGATGATATCAAGGCTTACCGCTCTTCAGAAATGAACTTAATTCGAATTAAGAGCTCTCATGGTACTACACTTGAAGTACCTGATCCCTGATGACCGCTGGTATTCTC	1080
V T E D D I K G L P C F Q N E T L I A I K A P H G T T T L E V P D P C E A G D Y L	305
CAGAGGAGATACAGAAATCGTATTAAAGAGTACCTCGGTCCTAATAGATGTTTACTTAGTTAGTCAATTTGACGATGGATTGAGAATTTGGTGGTGGTGGCAGACCTCAAGGCATACA	1200
Q R R Y R I V L K S T L G P I D V Y L V S Q F D D G F E N L G G A A T P P R H T	345
AATGTCCTCAAACTGGAGCTTGTGAACATTCATGCAACAAAGCTACACAAAGCAGCAAAATCAATCAATGGAATATAATATTCAGCAGCGCAGAAATCTCCACAAGATCTAGT	1320
M V P E P G P C E D L H A T N A T Q S S K S I N V E Y N I Q H R Q N T P Q D P S	385
TCTTCAAAATGATTAGGAGGATGACAAGGATAATCCCTTCAGATGTTAATACTGATGCTGATTACTGGCTCTAACAGAGGGGTGATGTTAGTATTGACATGTGGGAACAGCCACA	1440
S M D Y G G G H T R I I P S D V N T D A D Y M L T E G D V S I T D H W E T A P	425
GAAGTCAGTGGGACACGGCTGTGTTTTACCTGAAGATGTTAGCATCCACATGCACATCATAGTCCGCGGATGCAGGTTCCAAGCATGGATCAACCATTAAGGTCATGGCGGTGAAAC	1560
E V Q W D T A V F L P E D E D V S I P H A H E S P A R M Q V P S M D Q P	458
GCACATATGGAATTCCTGGAGTGTCTTTTCAGAAAACTGATTTCAAAATTGAAGAATGAGGCGACCAATTCAGACTGATCAGCTTGTGAATTTGCTGTTGTTATGCACAGGATT	1680
TTGGCAACTAATCTATCAGTCTGCTGCTGTTTGTCTGGACCTGTGCTCGAGTTGAAGGCGCGCCCATGTGCATATGCACTTGAATTCGGGCTGCTATGCACATTCGGTATCTG	1800
CTTATTTCTGAACTGAGTATATTGTTGGGCGCAATAGTGCCCTGTAAGCTCTCTGGGAATTAATCAAGATCTTTTGAGCAAAAACAGTAGGGAAGTCCCTGTTGTGACTCTTCA	1920
TTATATAAATCGAGTTTATACAAAGGGGTAAAAAAAAGGCTAAAAAAA	1975

SUBSTITUTE SHEET (RULE 26)

**FIG. 2**

RECTIFIED SHEET (RULE 91)
ISA/EP

3/5

```

YmE2F -----
HuE2F1 -----
HuE2F2 -----
HuE2F3 -----
HuE2F4 -----
HuE2F5 -----
OmE2F MSKFFVNVAP INNSNSSSSM TTTSSNTQRH QHQHQHYCGS GTTCHTMVAR ALNYDLHGCT TSINNMMNNIV IKNESVDLDY DHVLSSSDSN

-----MSG GSRPPAACKI :QSLRPPPVF STPSRPPFAS PDCYHRFHAP TTPSATSGGG IGSGGVGGDI CEGLVIRTQL
-----M ALAGAPAGGP CAPALEALLG AGALMLLDSS QIVTISAQD ASAPPAPTGP AAPAAGPCDP
-----MLQGPRALAS AAGOTPKVVP AMSPTLWPS GLSSPOLCPA TATYITPLYP QTAPPAAPG TCI.DATPHGP
GEGAAVVAAA AAASMDKPAL LASPGFAAAA AAAAAPGAYI QILTTNTSTT SCSSSLQSGA VAA..GPLLP SAPCAEQTAG SLLYTTPHGP
-----
SNGGVAHLR CHVYISLDKG HNTGAVATAA AAATAGCTQQ QLQQQHHHQN QQQRKATGXS NDITNYKVKX RRPFAVSDI HPKXQAKQSA

KRRATREENN AAESSDCMIV TTGTGNPLL TPVSGKAVKN SKSKTKNNKA GPQTPTPN.V GSPLNPSTPA GTORYDSSLG LLTKKFFVLL
DLLLFATPQA PRPTPSAPRP ALGRPPVKRR LDL.ETCHQY LAESSGPAGK RGR..... HPKGVKSPG EKSRVETSLN LLTKKFFVLL
EGQ..VVRCL PA..... GRLPKARR LDLEGIGRPV VPE.FPTPKG KCIKVDGL.P SPKTPKSPG EKTRYDTSLG LLTKKFFVLL
SSRAGLLOQP PALGRGGSGG G.GGPPAKRR LELGESGHQY LSDGLKTPKG KGRALRS.P DSPKTPKSPS EKTRYDTSLG LLTKKFFVLL
-----MAE.A G.PQAPPPPG TPSRHEKSLG LLTKKFFVLL
-----MAAEPAS SQQAPAGQG QGQKPPPPPP QAQAPQPP.P P.POLGGAGG GSSRHEKSLG LLTKKFFVLL
HHQTVYQKHT ASSAPQQLRH SHHQLRHDAO AELQEDVVER VAKPASHHFF SLSTPQQQLA ASVASSSSSU URNRAETSLG LLTKKFFVLL

KQAEQGLDL NAAAEHLV. QKRIYDITN VLEGILIER TLDRIERHMG LDD..SGVEL DNGLSCQTE VENNLQSOA LDERISDMRE
SHSADGVVDL NAAAEHLV. QKRIYDITN VLEGILIER KSNHICMLG SHITVGVG.. GRLEGLTQD LRQLQESBQQ LIDHLMNICTT
SESEQGVVDL NAAAEHLV. QKRIYDITN VLEGILIER KSNHICMLG RGMFEDPTR. FGKQOQLQOE LKEHMTEDA LDDLIOQCSL
SQSPQGVLDI NAAAEHLV. QKRIYDITN VLEGILIER KSNHICMLG CSLSDDCGN. AGQCOGLSKE VTSLSQEKK LDELIOQCSL
QEAQGVLDL KLAADILAVR QKRIYDITN VLEGILIER KSNHICMLG VPGONTREI ADKLIELKAE IEDQLHORE LDDHKKVWVQ
QEAQGVLDL KLAADILAVR QKRIYDITN VLEGILIER KSNHICMLG VGAGCNTKEV IDRLRYLKAE IEDQLHORE LDDHKKVWVQ
QESFDGVVDL NAAAEHLV. QKRIYDITN VLEGILIER KSNHICMLG GQSMVS.... QERSRHLEAD SLPLEQOE NEKALDLMRE

KLRLTEDEN SQ.RLAYVTI DDKGLPCFO NEHLIAKAP HGTLEVPDP GEAGDYLQRR YRMLRSLTG PIDVYLVSF DQGFENLQGA
QLRLTEDEN SQ.RLAYVTI DDKGLPCFO NEHLIAKAP HGTLEVPDP GEAGDYLQRR YRMLRSLTG PIDVYLVSF DQGFENLQGA
SFKHLDEKA NK.RLAYVTI DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK
DLKILTEDSE NK.RLAYVTI DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK
SIRNVLEDV NS.CLAYVTI DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK
SIRNVLEDV NN.RFSLVTI DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK
NLAEHSQVE NSGMAVVTI DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK DDYRAVGNFK

ATPPRHNTVP KPGPCEDLHA TNATQSSKSI NVEYNIQHRQ NTPQD..... PSS SNOYGGMTRI .....
PSQEVTS.E ENRATOSATI VSPPPSSPP..... SSIT TOPSQSLLSL EQEPLLSRMG
PLPSTSTLCP SPDSAQPSST TOPSIMEPTA SSVPA..... PART PQQAPPPPSL VP...LEAD
NNQOHNGNIP KP.ASKDL... ASTN SCHSDCSVSM GN...LSP:
VPPPEDL.LQ SPSAVSTPPP LPKPALAQSO EASRNSPQL TPTAVPGSAE VQCMAGPAE TVSGGPGTD SKDSGELSSL PLQPTTLDR
VPPPEDL.TQ PSSQSLTPVT PQKSSMATQN..... ISGD IDELMSDVF...PULLRLSPT
.PGAGYVGAR GAGCVRTATS TRLHPLTNR LNDPLFNNU AMSTKGLFQT PYRSARNLSK SIEEAAKQSQ PEYNNICDIA MSQHNNLQOQ

SLRA..... P VOEDRLSPLV AADSLLE... HVR EDF.SCLL... PEEFISLSPP
SL..... LE LPHPLLO... OTE DQFLSPTLAC SSPLISFSPG
AS..... PANLLO... QTE DO.IPSNLE.. GPFNVLIPP
PLOSSALLOS SSSSSSSSSS SSNSNSSSSS GPNPSTSEF IKADPTGVLE LPKELSEIFD PTRECHSSEL LEELMSSEVF..APLLRLSPP
ALQOTSATDI SSAGS..... PTKRRREAAA AATAAGSSTT ATTLNLSHNN
QQQQQQQLLQ QPEEDVDVE LNLVPTLTN PVVRTHQFQ HQGPSIQELF SSLETSSPPT PTKRRREAAA AATAAGSSTT ATTLNLSHNN

QFAAVAESNN GSSSGGGGGG NNSNNSNHS NHSSSNNSKS QPRTIGYGS ORRDVPMYN CAMEGATTS ATADTTAATS RSAAASSLQM

QFAAVAESNN GSSSGGGGGG GYGSIAGAG ANADPHQPS HDRNSLPPGV ADCDANSNS SVTLQGLDAL FNDIGSOYFS NDAFVSINP

NTP.DY.W.I GEGE.MIOM KETAFEVQWD TAVFLPEDVS IPHAHSPRM QVPSMDQP*
HEALDY.FEL EEEG.EIOL FD..CCFGDL TPL...DF*
LCCDDY.MCI EAGEG.ISOL FD.SVOLGDL.....LIN*
LLCEDY.MCI GEEEG..SOL FD.AIDLEKL PLVEDFNCS*
PCCDDY.MCI DESEG.VOOL FD.....VPVLNL*
PCCDDY.MCI DDNEG.VOOL FD.....VOILNY*
PCCDDY.MCI NANEG..DFL FD.GSEATGP*

```

FIG. 3

SUBSTITUTE SHEET (RULE 26)

4/5

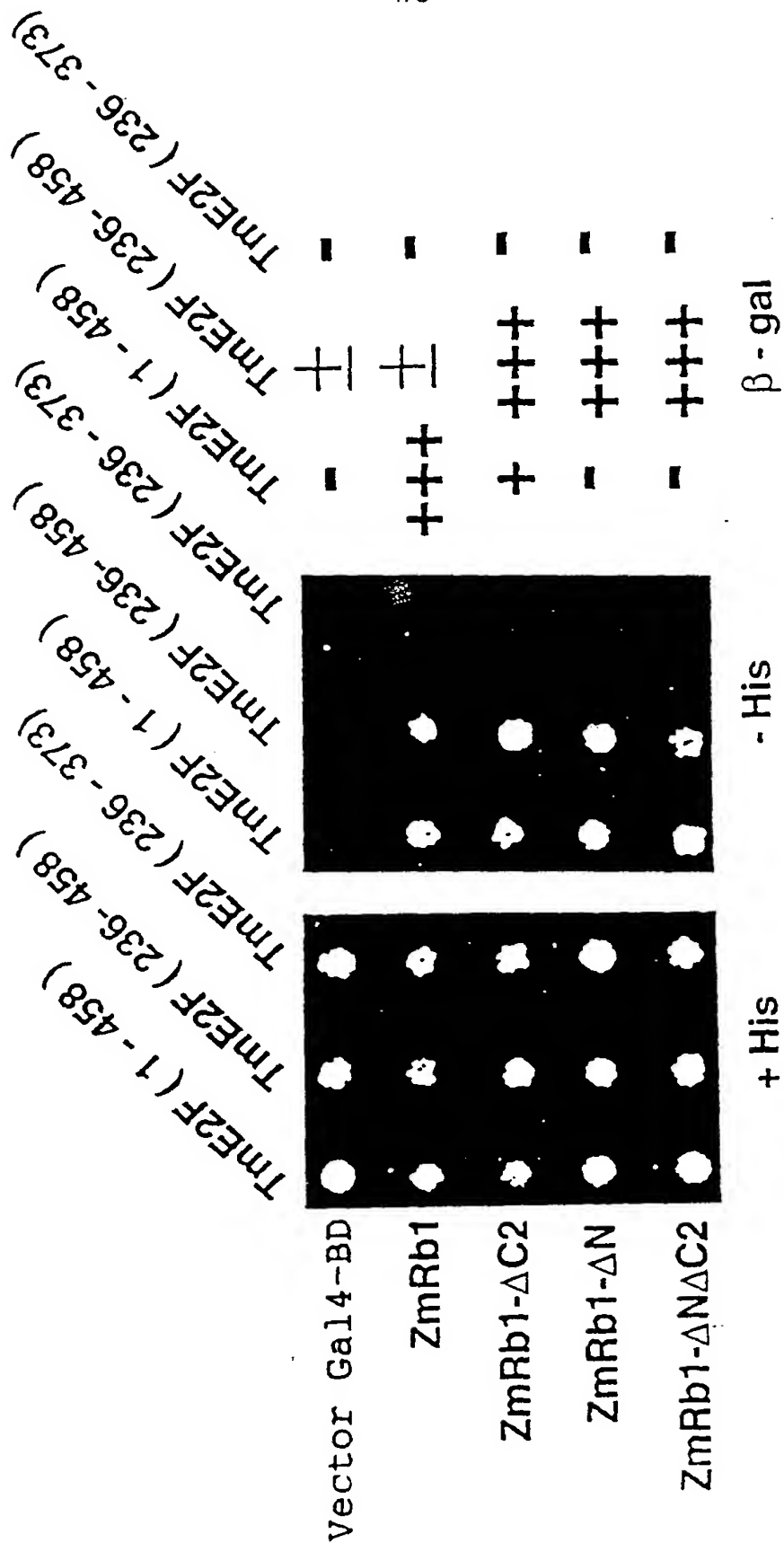
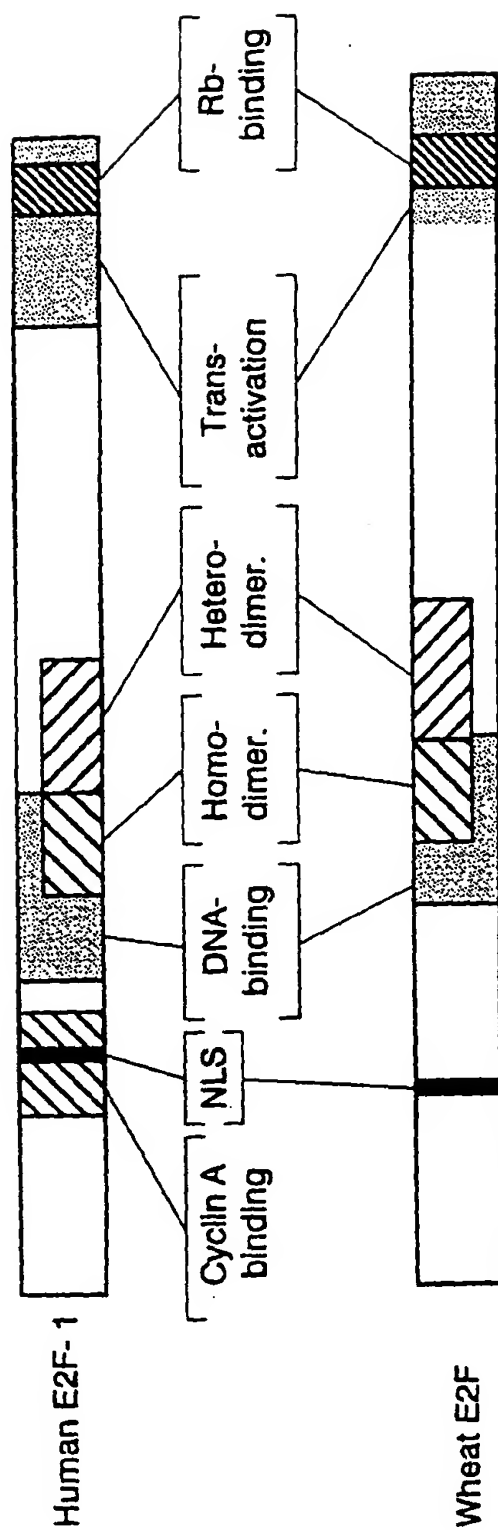


FIG. 4

SUBSTITUTE SHEET (RULE 26)

**FIG. 5**

RECTIFIED SHEET (RULE 91)
ISA/EP

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- 5 (i) APPLICANT:
(A) NAME: CONSEJO SUPERIOR DE INVESTIGACIONES
CIENTIFICAS
(B) STREET: SERRANO 117
10 (C) CITY: MADRID
(E) COUNTRY: SPAIN
(F) POSTAL CODE (ZIP): E-28006

(A) NAME: CRISANTO GUTIERREZ-ARMENTA
15 (B) STREET: CSIC-UAM UNIVERSIDAD AUTONOMA CANTOBLANCO
(C) CITY: MADRID
(E) COUNTRY: SPAIN
(F) POSTAL CODE (ZIP): 28049

20 (A) NAME: ELENA RAMIREZ-PARRA
(B) STREET: CIC-UAM UNIVERSIDAD AUTONOMA CANTOBLANCO
(C) CITY: MADRID
(E) COUNTRY: SPAIN
(F) POSTAL CODE (ZIP): 28049
25 (A) NAME: QI XIE
(B) STREET: CSIC-UAM UNIVERSIDAD AUTONOMA CANTOBLANCO
(C) CITY: MADRID
(E) COUNTRY: SPAIN
30 (F) POSTAL CODE (ZIP): 28049

(ii) TITLE OF INVENTION: TRANSGENIC PLANT CELLS

35 (iii) NUMBER OF SEQUENCES: 7

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
40 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

- 45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

55 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Triticum monococcum*

5 (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..48

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TAC TGG CTC CTA ACA GAG GGT GAT GTT AGT ATT ACT GAC ATG TGG GAA
48
Tyr Trp Leu Leu Thr Glu Gly Asp Val Ser Ile Thr Asp Met Trp Glu
15 1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 2:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Tyr Trp Leu Leu Thr Glu Gly Asp Val Ser Ile Thr Asp Met Trp Glu
1 5 10 15

30 (2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
35 (A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45 (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Triticum monococcum*

(ix) FEATURE:
50 (A) NAME/KEY: CDS
(B) LOCATION:1..48

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

55

TACNNNNNNN NNNNNNNNNN NGATNNNNNN NNNNNNGACA TGTGGGAA
48

(2) INFORMATION FOR SEQ ID NO: 4:

5

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

10

- (ii) MOLECULE TYPE: DNA (genomic)

15

- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Triticum monococcum*

20

- (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:1..48

25

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

30

Tyr Xaa Xaa Xaa Xaa Xaa Xaa Asp Xaa Xaa Xaa Xaa Asp Met Trp Glu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 5:

35

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1974 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

40

- (ii) MOLECULE TYPE: DNA (genomic)

- (iii) HYPOTHETICAL: NO

45

- (iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Triticum monococcum*

50

- (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:166..1539

55

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AATTCGGGCAC GAGCCACCCC ACCTACCTCC CGCCGCGGCC GCCGCCACGG GAACCCCTATC
 60
 5 TCCGGCGAGC CCCCCCGAT GCCCTCCTGC CTTCTCTGAA GCCGAAGACG CCCATCGCTC
 120
 CCGGGAGTCG GGGGTCCCGC AGCGCGCGAT CGCGAGATCG GGCTT ATG TCT GGG
 174
 10 Met Ser Gly
 GGC GGC AGG CCG CCG GCT GCG CAA AAA ATC CTG CAG TCT CTG CGC CCG
 222
 15 Gly Gly Arg Pro Pro Ala Ala Gln Lys Ile Leu Gln Ser Leu Arg Pro
 20 25 30 35
 CCC CCG GTG TTC TCC ACG CCG TCG CGG CCT CCC TTC GCC TCA CCC GAC
 270
 20 Pro Pro Val Phe Ser Thr Pro Ser Arg Pro Pro Phe Ala Ser Pro Asp
 40 45 50
 GAC TAC CAC CGC TTT CAT GCG CCG ACT ACC CCT TCT GCC ACT GGC TCC
 318
 25 Asp Tyr His Arg Phe His Ala Pro Thr Thr Pro Ser Ala Thr Gly Ser
 55 60 65
 GGC GGC ATC GGC TCC GGT GGT GTT GGC GGC GAT ATT GAT GAG GGG CTT
 366
 30 Gly Gly Ile Gly Ser Gly Gly Val Gly Gly Asp Ile Asp Glu Gly Leu
 70 75 80
 GTT ATC CGG ACG CAG CTA AAA AGA AAA GCC ACA CGC GAA GAA AAT AAT
 414
 35 Val Ile Arg Thr Gln Leu Lys Arg Lys Ala Thr Arg Glu Glu Asn Asn
 85 90 95
 CCG GCT GAG TCG AGT GAC TGT ATG ATT GTC ACC ACT GGA GTT ACT GGC
 462
 40 Ala Ala Glu Ser Ser Asp Cys Met Ile Val Thr Thr Gly Val Thr Gly
 100 105 110 115
 AAT CCG CTA CTC ACC CCA GTG TCT GGA AAA GCT GTT AAG AAT TCT AAA
 510
 45 Asn Pro Leu Leu Thr Pro Val Ser Gly Lys Ala Val Lys Asn Ser Lys
 120 125 130
 TCA AAG ACT AAG AAC AAT AAA GCT GGG CCT CAG ACA CCT ACG CCA AAT
 558
 50 Ser Lys Thr Lys Asn Asn Lys Ala Gly Pro Gln Thr Pro Thr Pro Asn
 135 140 145
 GTT GGC TCA CCA CTC AAT CCA TCA ACT CCT GCT GGT ACT TGC CGC TAT
 606
 55 Val Gly Ser Pro Leu Asn Pro Ser Thr Pro Ala Gly Thr Cys Arg Tyr

	150	155	160
	GAC AGT TCG TTA GGA CTT CTG ACA AAG AAG TTC ATC AAT TTG CTG AAG		
	654		
5	Asp Ser Ser Leu Gly Leu Leu Thr Lys Lys Phe Ile Asn Leu Leu Lys		
	165	170	175
	CAA GCT GAG GAT GGC ATT CTA GAT TTG AAT AAT GCT GCA GAA ACA CTA		
	702		
10	Gln Ala Glu Asp Gly Ile Leu Asp Leu Asn Asn Ala Ala Glu Thr Leu		
	180	185	190 195
	GAG GTT CAA AAG CGA CGC ATA TAT GAC ATC ACA AAT GTC CTC GAA GGA		
	750		
15	Glu Val Gln Lys Arg Arg Ile Tyr Asp Ile Thr Asn Val Leu Glu Gly		
		200	205 210
	ATT GGT CTT ATA GAA AAG ACA CTT AAA AAC AGA ATT CGT TGG AAG GGC		
	798		
20	Ile Gly Leu Ile Glu Lys Thr Leu Lys Asn Arg Ile Arg Trp Lys Gly		
		215	220 225
	TTG GAT GAT TCA GGA GTG GAA TTA GAT AAT GGC CTT TCA GGT TTG CAG		
	846		
25	Leu Asp Asp Ser Gly Val Glu Leu Asp Asn Gly Leu Ser Gly Leu Gln		
		230 235	240
	ACA GAA GTT GAA AAT CTT AAT TTG CAG GAG CAA GCC TTA GAT GAG CGT		
	894		
30	Thr Glu Val Glu Asn Leu Asn Leu Gln Glu Gln Ala Leu Asp Glu Arg		
		245 250	255
	ATA AGT GAT ATG CGC GAA AAG CTA AGG GGG TTA ACG GAA GAT GAG AAC		
	942		
35	Ile Ser Asp Met Arg Glu Lys Leu Arg Gly Leu Thr Glu Asp Glu Asn		
		260 265	270 275
	AGT CAA AGA TGG CTC TAT GTG ACG GAA GAT GAT ATC AAG GGA TTA CCC		
	990		
40	Ser Gln Arg Trp Leu Tyr Val Thr Glu Asp Asp Ile Lys Gly Leu Pro		
		280 285	290
	TGC TTT CAG AAT GAA ACT CTA ATT GCA ATA AAA GCT CCT CAT GGT ACT		
	1038		
45	Cys Phe Gln Asn Glu Thr Leu Ile Ala Ile Lys Ala Pro His Gly Thr		
		295 300	305
	ACA CTT GAA GTA CCT GAT CCT GAT GAG GCT GGT GAT TAT CTC CAG AGG		
	1086		
50	Thr Leu Glu Val Pro Asp Pro Asp Glu Ala Gly Asp Tyr Leu Gln Arg		
		310 315	320
	AGA TAC ACA ATC GTA TTA AGA AGT ACC CTG GGT CCA ATA GAT GTT TAC		
	1134		
55	Arg Tyr Arg Ile Val Leu Arg Ser Thr Leu Gly Pro Ile Asp Val Tyr		

	325	330	335
	TTA GTT AGT CAA TTT GAC GAT GGA TTT GAG AAT TTG GGT GGT GCT GCG		
	1182		
5	Leu Val Ser Gln Phe Asp Asp Gly Phe Glu Asn Leu Gly Gly Ala Ala		
	340	345	350 355
	ACA CCT CCA AGG CAT ACA AAT GTC CCA AAA CCT GGA CCT TGT GAA GAC		
	1230		
10	Thr Pro Pro Arg His Thr Asn Val Pro Lys Pro Gly Pro Cys Glu Asp		
	360	365	370
	TTA CAT GCA ACA AAC GCT ACA CAA AGC AGC AAA TCA ATC AAT GTG GAA		
	1278		
15	Leu His Ala Thr Asn Ala Thr Gln Ser Ser Lys Ser Ile Asn Val Glu		
	375	380	385
	TAT AAT ATT CAG CAC AGG CAG AAT ACT CCA CAA GAT CCT AGT TCT TCA		
	1326		
20	Tyr Asn Ile Gln His Arg Gln Asn Thr Pro Gln Asp Pro Ser Ser Ser		
	390	395	400
	AAT GAT TAT GGA GGG ATG ACA AGG ATA ATC CCT TCA GAT GTT AAT ACT		
	1374		
25	Asn Asp Tyr Gly Gly Met Thr Arg Ile Ile Pro Ser Asp Val Asn Thr		
	405	410	415
	GAT GCT GAT TAC TGG CTC CTA ACA GAG GGT GAT GTT AGT ATT ACT GAC		
	1422		
30	Asp Ala Asp Tyr Trp Leu Leu Thr Glu Gly Asp Val Ser Ile Thr Asp		
	420	425	430 435
	ATG TGG GAA ACA GCA CCA GAA GTG CAG TGG GAC ACC GCT GTG TTT TTA		
	1470		
35	Met Trp Glu Thr Ala Pro Glu Val Gln Trp Asp Thr Ala Val Phe Leu		
	440	445	450
	CCT GAA GAT GTT AGC ATC CCA CAT GCA CAT CAT AGT CCG CGG ATG CAG		
	1518		
40	Pro Glu Asp Val Ser Ile Pro His Ala His His Ser Pro Arg Met Gln		
	455	460	465
	GTT CCA AGC ATG GAT CAA CCA TAAGGTCATG GCGGTGAAAA CTTGACATAT		
	1569		
45	Val Pro Ser Met Asp Gln Pro		
	470		
	GGAATTCCTG GAGTGCTGTT TCAGAAAATA CTGATTTCAA AATGGAAAGA TCAGGGCAGC		
	1629		
50	AAGTTCAGAC TGATCACCGT TCTGAATTTG CTGTTTGTTA TGGACACGAT TGGTGCCAAC		
	1689		
	TAACTTATCA GTCTGCTGCC TTGTTTGTTT TGGCACCTGT CCTTCAGTTG AAAAGGCGCC		
55	1749		

CATGTGCATA TTGCACCTTG AATTCGGGCT GCTATGCACA TTCGGTATCT GCTTTATTTTC
1809

5 TCTAACTGAG TATATTTTGC AAGGCAATAG TGGCTCTGTA GCTCTCTTGG GAATTAATAC
1869

GAATCTTTTT GAGCAAAAAC AGTAGGGAAG TCCCCTGTTG TGACTCTTTC ATTATATAAA
1929

10 TGGAGTTTAT ACAAAGGGGT AAAAAAAAAA AAAAAAAAAA AAAAA
1974

15 (2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 458 amino acids
(B) TYPE: amino acid
20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

25 Met Ser Gly Gly Gly Arg Pro Pro Ala Ala Gln Lys Ile Leu Gln Ser
1 5 10 15

Leu Arg Pro Pro Pro Val Phe Ser Thr Pro Ser Arg Pro Pro Phe Ala
20 25 30

30 Ser Pro Asp Asp Tyr His Arg Phe His Ala Pro Thr Thr Pro Ser Ala
35 40 45

35 Thr Gly Ser Gly Gly Ile Gly Ser Gly Gly Val Gly Gly Asp Ile Asp
50 55 60

Glu Gly Leu Val Ile Arg Thr Gln Leu Lys Arg Lys Ala Thr Arg Glu
65 70 75 80

40 Glu Asn Asn Ala Ala Glu Ser Ser Asp Cys Met Ile Val Thr Thr Gly
85 90 95

Val Thr Gly Asn Pro Leu Leu Thr Pro Val Ser Gly Lys Ala Val Lys
100 105 110

45 Asn Ser Lys Ser Lys Thr Lys Asn Asn Lys Ala Gly Pro Gln Thr Pro
115 120 125

50 Thr Pro Asn Val Gly Ser Pro Leu Asn Pro Ser Thr Pro Ala Gly Thr
130 135 140

Cys Arg Tyr Asp Ser Ser Leu Gly Leu Leu Thr Lys Lys Phe Ile Asn
145 150 155 160

55 Leu Leu Lys Gln Ala Glu Asp Gly Ile Leu Asp Leu Asn Asn Ala Ala

	165	170	175
	Glu Thr Leu Glu Val Gln Lys Arg Arg Ile Tyr Asp Ile Thr Asn Val		
	180	185	190
5	Leu Glu Gly Ile Gly Leu Ile Glu Lys Thr Leu Lys Asn Arg Ile Arg		
	195	200	205
10	Trp Lys Gly Leu Asp Asp Ser Gly Val Glu Leu Asp Asn Gly Leu Ser		
	210	215	220
	Gly Leu Gln Thr Glu Val Glu Asn Leu Asn Leu Gln Glu Gln Ala Leu		
	225	230	235
15	Asp Glu Arg Ile Ser Asp Met Arg Glu Lys Leu Arg Gly Leu Thr Glu		
	245	250	255
	Asp Glu Asn Ser Gln Arg Trp Leu Tyr Val Thr Glu Asp Asp Ile Lys		
	260	265	270
20	Gly Leu Pro Cys Phe Gln Asn Glu Thr Leu Ile Ala Ile Lys Ala Pro		
	275	280	285
25	His Gly Thr Thr Leu Glu Val Pro Asp Pro Asp Glu Ala Gly Asp Tyr		
	290	295	300
	Leu Gln Arg Arg Tyr Arg Ile Val Leu Arg Ser Thr Leu Gly Pro Ile		
	305	310	315
30	Asp Val Tyr Leu Val Ser Gln Phe Asp Asp Gly Phe Glu Asn Leu Gly		
	325	330	335
	Gly Ala Ala Thr Pro Pro Arg His Thr Asn Val Pro Lys Pro Gly Pro		
	340	345	350
35	Cys Glu Asp Leu His Ala Thr Asn Ala Thr Gln Ser Ser Lys Ser Ile		
	355	360	365
40	Asn Val Glu Tyr Asn Ile Gln His Arg Gln Asn Thr Pro Gln Asp Pro		
	370	375	380
	Ser Ser Ser Asn Asp Tyr Gly Gly Met Thr Arg Ile Ile Pro Ser Asp		
	385	390	395
45	Val Asn Thr Asp Ala Asp Tyr Trp Leu Leu Thr Glu Gly Asp Val Ser		
	405	410	415
	Ile Thr Asp Met Trp Glu Thr Ala Pro Glu Val Gln Trp Asp Thr Ala		
	420	425	430
50	Val Phe Leu Pro Glu Asp Val Ser Ile Pro His Ala His His Ser Pro		
	435	440	445
55	Arg Met Gln Val Pro Ser Met Asp Gln Pro		
	450	455	

(2) INFORMATION FOR SEQ ID NO: 7:

5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
20 (A) ORGANISM: *Triticum monococcum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

25 Arg Thr Gln Leu Lys Arg Lys Ala Thr Arg Glu Glu
1 5 10